

BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 22 MARS 2005

Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

Martine PLANCHE

DOCUMENT DE
PRIORITÉ

PRÉSENTÉ OU TRANSMIS
CONFORMÉMENT À LA RÈGLE
17.1. a) OU b)

INSTITUT
NATIONAL DE
LA PROPRIÉTÉ
INDUSTRIELLE

SIEGE
26 bis, rue de Saint-Petersbourg
75800 PARIS cedex 08
Téléphone : 33 (0)1 53 04 53 04
Télécopie : 33 (0)1 53 04 45 23
www.inpi.fr





BREVET D'INVENTION

CERTIFICAT D'UTILITE

26bis, rue de Saint-Petersbourg
75800 Paris Cédex 08
Téléphone: 01 53.04.53.04 Télécopie: 01.42.94.86.54

Code de la propriété intellectuelle-livre VI

REQUÊTE EN DÉLIVRANCE

DATE DE REMISE DES PIÈCES: N° D'ENREGISTREMENT NATIONAL: DÉPARTEMENT DE DÉPÔT: DATE DE DÉPÔT:	Brigitte RUELLAN-LEMONNIER THOMSON 46, quai Alphonse Le Gallo 92648 BOULOGNE cedex France
Vos références pour ce dossier: PF040033	

1 NATURE DE LA DEMANDE	
Demande de brevet	
2 TITRE DE L'INVENTION	
	PROCEDE DE CODAGE D'UNE SEQUENCE D'IMAGES PAR ANALYSE TEMPORELLE HIERARCHIQUE
3 DECLARATION DE PRIORITE OU REQUETE DU BENEFICE DE LA DATE DE DEPOT D'UNE DEMANDE ANTERIEURE FRANCAISE	Pays ou organisation Date N°
4-1 DEMANDEUR	
Nom Suivi par Rue Code postal et ville Pays Nationalité Forme juridique N° SIREN Code APE-NAF N° de téléphone N° de télécopie	THOMSON LICENSING S.A. RUELLAN Brigitte 46, quai Alphonse Le Gallo 92100 Boulogne France France Société anonyme 383 461 191 322A + 33 1 41 86 50 00 + 33 1 41 8656 33
5A MANDATAIRE	
Nom Prénom Qualité Cabinet ou Société Rue Code postal et ville N° de téléphone N° de télécopie Courrier électronique	RUELLAN-LEMONNIER Brigitte Mandataire professionnel communautaire, Pouvoir général: 12778 THOMSON 46, quai Alphonse Le Gallo 92648 BOULOGNE cedex + 33 1 41 86 52 77 + 33 1 41 86 56 33 brigitte.ruellan@thomson.net

6 DOCUMENTS ET FICHIERS JOINTS		Fichier électronique	Pages	Détails
Texte du brevet	textebrevet.pdf	20		D 15, R 4, AB 1
Dessins	dessins.pdf	3		page 3, figures 5
Désignation d'inventeurs				
Pouvoir général				
7 MODE DE PAIEMENT				
Mode de paiement	Prélèvement du compte courant			
Numéro du compte client	3334			
8 RAPPORT DE RECHERCHE				
Etablissement immédiat				
9 REDEVANCES JOINTES				
	Devise	Taux	Quantité	Montant à payer
062 Dépôt	EURO	0.00	1.00	0.00
063 Rapport de recherche (R.R.)	EURO	320.00	1.00	320.00
068 Revendication à partir de la 11ème	EURO	15.00	2.00	30.00
Total à acquitter	EURO			350.00

La loi n°78-17 du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire. Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI.

Signé par

Signataire: FR, THOMSON, B.Ruellan

Emetteur du certificat: DE, D-Trust GmbH, D-Trust for EPO 2.0

Fonction

Mandataire agréé (Mandataire 1)



BREVET D'INVENTION CERTIFICAT D'UTILITÉ

Réception électronique d'une soumission

Il est certifié par la présente qu'une demande de brevet (ou de certificat d'utilité) a été reçue par le biais du dépôt électronique sécurisé de l'INPI. Après réception, un numéro d'enregistrement et une date de réception ont été attribués automatiquement.

Demande de brevet : X

Demande de CU :

DATE DE RECEPTION	2 mars 2004	
TYPE DE DEPOT	INPI (PARIS) - Dépôt électronique	Dépôt en ligne: X
N° D'ENREGISTREMENT NATIONAL ATTRIBUE PAR L'INPI	0450420	Dépôt sur support CD:
Vos références pour ce dossier	PF040033	

DEMANDEUR

Nom ou dénomination sociale	THOMSON LICENSING S.A.
Nombre de demandeur(s)	1
Pays	FR

TITRE DE L'INVENTION

PROCEDE DE CODAGE D'UNE SEQUENCE D'IMAGES PAR ANALYSE TEMPORELLE HIERARCHIQUE

DOCUMENTS ENVOYES

package-data.xml	Requetefr.PDF	application-body.xml
Design.PDF	ValidLog.PDF	fee-sheet.xml
FR-office-specific-info.xml	Comment.PDF	textebrevet.pdf
dessins.pdf	indication-bio-deposit.xml	request.xml

EFFECTUE PAR

Effectué par:	B.Ruellan
Date et heure de réception électronique:	2 mars 2004 17:29:45
Empreinte officielle du dépôt	74:DB:38:3F:A3:3A:04:F1:D5:2D:D6:AD:A7:C5:A4:4C:1A:21:61:32

/ INPI PARIS, Section Dépôt /

SIEGE SOCIAL
INSTITUT 28 bis, rue de SaintPetersbourg
NATIONAL DE 75000 PARIS cedex 08
LA PROPRIETE Téléphone : 01 53 04 53 04
INDUSTRIELLE Télécopie : 01 42 93 53 90

L'invention concerne un procédé de codage et de décodage vidéo d'une séquence d'image codée par analyse temporelle hiérarchique exploitant le filtrage temporel compensé en mouvement.

Le domaine est celui de la compression vidéo basée sur des schémas à échelonnabilité spatiale et/ou temporelle dits aussi « scalables ». Il s'agit par exemple d'un codage en ondelettes 2D+t comportant un filtrage temporel compensé en mouvement.

Dans les standards de codage dits hybrides, tels que MPEG-1, MPEG-2, MPEG-4, h264, comme dans la plupart des schémas de codage sous-bande 2D+t, tels que MC-EZBC acronyme de l'anglais Motion Compensated Embedded Zero Block Context, la première étape de la chaîne de codage consiste à tirer parti de la redondance temporelle entre images successives, avant d'exploiter la redondance spatiale au sein d'une image.

La figure 1 représente un schéma d'un codeur vidéo selon l'art antérieur.

Le signal vidéo est transmis à un circuit d'analyse temporelle 1. Un circuit d'estimation de mouvement 2 est relié à ce premier circuit pour estimer le mouvement entre deux images reçues par le codeur. Les informations de mouvement sont transmises au circuit 1 et à un circuit 6 de codage, par exemple sous forme de champs de vecteurs mouvement. La sortie du circuit 1 est transmise à un circuit d'analyse spatiale 3 qui extrait, de la texture, les coefficients de fréquence de l'image. Ces coefficients sont ensuite quantifiés puis codés par un codage entropique, circuit 4. Ces informations codées et celles de mouvement sont transmises à un circuit de mise en paquets ou paquetiseur 5 qui envoie les données vidéo sous forme de paquets vidéo pour constituer le flux de données vidéo.

Le circuit d'analyse temporelle 1 réalise une prédiction temporelle compensée en mouvement dans le cas d'un schéma hybride ou un filtrage temporel compensé en mouvement dans le cas d'un schéma de codage sous-bande. Les algorithmes de codage avec prédiction temporelle consistent à opérer de la compensation de mouvement pour générer des images de prédiction qui vont servir ultérieurement dans le processus de codage. Ces algorithmes reposent sur le même principe. Les images à coder sont prédites à partir d'une ou plusieurs images précédemment codées,

appelées images de référence. C'est le cas dans les standards MPEG vidéo avec les images Prédites (P) et Bi-directionnelles ou Bi-prédites (B). La prédiction consiste à effectuer une compensation de mouvement à partir de ces images de référence et de vecteurs mouvement associés à l'image courante. Ce qui est ensuite codé est le résidu de la prédiction, c'est-à-dire la différence entre l'image courante et l'image de prédiction temporelle. Le mouvement est par exemple décrit par blocs de pixels et la compensation en mouvement effectuée par bloc.

Le circuit d'analyse spatiale 3 réalise une décomposition en ondelettes ou une transformée cosinus discrète. Le codage entropique du circuit 4 peut être un codage de type VLC, acronyme de l'anglais Variable Length Coding, ou un codage de type arithmétique.

Le circuit de paquetisation a pour fonction de découper l'information de texture et de mouvement provenant respectivement du circuit de codage entropique et du circuit de codage des champs de mouvement en sous-ensembles cohérents selon leur fréquence spatiale, temporelle, et leur importance, par exemple leur poids dans une approche de codage par plans de bits. Ainsi le flux binaire obtenu est échelonnable ou "scalable" indépendamment en résolution, en fréquence trame et en fidélité.

Les champs de mouvement estimés correspondent à la résolution de la source. L'étape de compensation en mouvement du codeur, qu'elle soit faite par filtrage ou prédiction est donc exécutée sur des images à pleine résolution tandis qu'au décodeur, l'étape de compensation de mouvement peut être exécutée sur des images de résolution moindre, avec des champs de mouvement remis à l'échelle.

Le but principal du filtrage temporel avec compensation en mouvement, aussi connu sous l'acronyme MCTF, de l'appellation anglaise Motion Compensated Temporal Filtering, est de générer des images haute fréquence H nécessitant le minimum de bits pour le codage. Une autre contrainte est de réduire autant que possible le nombre de pixels non connectés, c'est à dire non reliés par les vecteurs mouvement. Ces deux aspects sont directement dépendant de la qualité du champ de mouvement. Les informations filtrées sont d'autant plus corrélées que le mouvement exploité pour le filtrage est de bonne qualité.

Dans certains cas, cette qualité de mouvement n'est pas satisfaisante, par exemple si le mouvement est trop complexe, trop important ou s'il y a une coupure de scène. Cela résulte alors en un excès d'énergie dans les images hautes fréquences H et dans un nombre important de pixels non connectés dans les images basse fréquence L, en final en un coût de codage important et une mauvaise répartition de la qualité visuelle au sein du groupe d'images décodées.

L'exploitation de GOP adaptatif permet de prendre en compte ce problème. La taille du GOP est adaptée de manière dynamique en fonction du pourcentage de pixels non connectés. Si le nombre total de pixels non connectés dans une image est supérieur à un seuil prédéfini, par exemple de l'ordre de 60 à 75 %, lors de l'estimation de mouvement effectuée à un niveau de décomposition I, la décomposition temporelle en sous-bandes est arrêtée. Le GOP initial est modifié en taille pour donner deux GOPs, la taille du premier GOP étant automatiquement forcée à 2^1 . Cette solution oblige cependant à avoir des GOPs de taille variable, la gestion de tels GOPs est complexe.

Une autre solution consiste à exploiter d'autres modes de codage tels que le codage en mode intra, lors de la décomposition temporelle. S'il existe un nombre de pixels non connectés important, un codage intra est réalisé. Mais la dynamique des images doit alors être adaptée à celle des images obtenues par filtrage MCTF, entraînant un surcoût de codage. Ceci au détriment des images codées selon le mode MCTF et donc de la qualité globale de la séquence d'images reconstituée.

L'invention a pour but de pallier les inconvénients précités.

Un des objets de l'invention est un procédé de codage d'une séquence d'images comprenant une analyse temporelle hiérarchique d'un groupe d'images réalisant un filtrage temporel compensé en mouvement de paires successives d'images pour fournir des images de fréquence temporelle basse et des images de fréquence temporelle haute à différents niveaux de décomposition temporelle, cette analyse réalisant, pour un niveau de décomposition temporelle donné et pour une paire d'images de fréquence temporelle basse, une étape d'estimation de mouvement d'une image courante B vers une image précédente de référence A pour fournir des vecteurs mouvement puis un filtrage temporel compensé en mouvement de

ces images pour fournir une image de fréquence temporelle basse (L) et une image de fréquence temporelle haute (H) d'un niveau de décomposition supérieur, le dit filtrage temporel étant remplacé par un codage en mode intra pour obtenir au moins une image de fréquence basse (L) ou haute (H) si
5 l'image courante a un niveau de corrélation avec une image précédente inférieur à un seuil, les images obtenues de fréquence basse (L) étant alors mises à l'échelle pour être adaptées, au niveau énergétique, aux images obtenues par le dit filtrage temporel compensé en mouvement, caractérisé en ce que, parmi l'image basse fréquence et les images haute fréquence
10 finales décomposées obtenues en fin d'analyse :

- il sélectionne les images obtenues par un codage intra d'une image du niveau de décomposition inférieur avec la condition supplémentaire, pour les images haute fréquence, que cette image provient elle-même d'un codage intra.

- 15 - il calibre l'image sélectionnée en effectuant au moins une étape inverse de l'étape de mise à l'échelle.

Le nombre d'étape inverses effectuées correspond au nombre de codages intra successifs d'une image basse fréquence (L) pour arriver à l'image sélectionnée s'il s'agit d'une image sélectionnée haute fréquence, ce
20 nombre étant incrémenté de un s'il s'agit de l'image sélectionnée basse fréquence (L).

Selon une mise en œuvre particulière, le procédé comporte, pour le calcul d'une image fréquence basse L ou haute H, un filtrage temporel entre l'image courante et une image suivante de la paire d'images suivante,
25 si la corrélation entre l'image courante et l'image précédente est inférieure à un seuil et si la corrélation entre l'image courante et cette image suivante est supérieure à un seuil, l'autre image H ou L étant obtenue par codage intra et en ce que cette opération de filtrage est assimilée au codage intra et non pas au filtrage temporel pour l'étape de sélection.

- 30 Selon une mise en œuvre particulière, le procédé affecte à chaque image du groupe d'images un numéro d'image, il effectue un suivi de ces images numérotées lors de la décomposition en attribuant un compteur pour chaque numéro, ce compteur étant actualisé à chaque étape,

- le compteur est incrémenté chaque fois qu'une image de fréquence basse (L) est obtenue en mode intra,

- le compteur est inchangé chaque fois qu'une image de fréquence haute (H) est obtenue en mode intra ou lors d'un filtrage temporel avec une image suivante,

- le compteur est remis à zéro chaque fois qu'une image est obtenue par filtrage temporel compensé en mouvement, avec une image précédente.

L'invention concerne également un procédé de décodage d'une séquence d'images codées selon le procédé précédemment décrit, caractérisé en ce qu'il effectue une étape de calibrage inverse d'images sélectionnées à décoder, la sélection des images et le nombre d'étapes inverses étant fonction d'une information associée à l'image à décoder. Cette information est par exemple la valeur du compteur affecté à l'image lors du codage.

L'invention concerne également un codeur pour la mise en œuvre du procédé décrit, comportant un circuit d'analyse temporelle exploitant le filtrage temporel compensé en mouvement et le codage intra, caractérisé en ce que le circuit sélectionne, parmi l'image basse fréquence et les images haute fréquence finales décomposées obtenues en fin d'analyse, les images obtenues par un codage intra d'une image du niveau de décomposition inférieur ou par un filtrage temporel entre l'image courante et une image suivante du niveau de décomposition inférieur, avec la condition supplémentaire, pour les images haute fréquence, que cette image provient elle-même d'un codage intra et en ce qu'il effectue au moins une étape inverse de l'étape de mise à l'échelle pour les images sélectionnées.

L'invention concerne également un décodeur pour le décodage de séquence d'images codées selon le procédé de la revendication 1, comportant un circuit de synthèse temporelle, caractérisé en ce que le circuit comporte des moyens pour effectuer un calibrage inverse d'images à décoder, la sélection des images et le nombre de calibrages inverses étant fonction d'une information associée à l'image à décoder et reçue par le décodeur.

Grâce à cette opération inverse de mise à l'échelle des images de type intra, les images codées selon le mode MCTF ne sont pas pénalisées. La répartition de la qualité des images sur l'ensemble d'un GOP comportant à la fois des images codées en intra ou selon le mode MCTF est améliorée. Les informations de mode de codage transmises au décodeur permettent de recalibrer les images intra pour effectuer la synthèse des images.

Le filtrage temporel étant réalisé de manière conditionnelle, selon la qualité du mouvement, la décomposition temporelle en sous-bandes peut être effectuée jusqu'au dernier niveau. Quelque soit la qualité du mouvement, la structure du GOP est conservée. La taille des GOPs peut être maintenue constante, facilitant la gestion de débit et de traitement, même si une coupure de scène a lieu au milieu d'un GOP. Si l'on exploite des GOPs de taille variable, la qualité d'image est améliorée.

D'autres particularités et avantages de l'invention ressortiront mieux dans la description suivante donnée à titre d'exemple non limitatif et faite en regard des figures annexées qui représentent :

- la figure 1 un schéma de codage selon l'art antérieur,
- la figure 2, un filtrage temporel compensé en mouvement, sur un GOP de 16 images,
- la figure 3, un circuit de codage,
- la figure 4, un organigramme de codage,
- la figure 5, un circuit de décodage.

La figure 2 représente de manière synthétique les opérations de filtrage temporel compensé en mouvement réalisées par le circuit d'analyse temporelle 4, avec une décomposition à 4 niveaux pour des GOPs comportant, dans cet exemple, 16 images représentées en traits gras.

Le mode de filtrage utilisé est appelé « lifting ». Au lieu d'utiliser un filtrage complexe pour le codage en ondelettes, exploitant un filtre linéaire de longueur importante, dans notre exemple le filtrage serait effectué sur un groupe de 16 images, ce procédé de filtrage consiste, de manière connue, à « factoriser » le filtre en exploitant des filtres de longueur limitée, par exemple deux si l'on choisit de filtrer les échantillons deux à deux, ce filtrage étant reconduit pour chaque niveau de décomposition. On considère donc le

cas où le filtrage dans le sens du mouvement est effectué sur des paires d'images. Le filtrage basse fréquence et haute fréquence, sur chacune des paires du GOP, produit respectivement, au premier niveau de décomposition temporelle, 8 images de fréquence temporelle basse (t-L) et 8 images de fréquence temporelle haute (t-H).

Les images de fréquence temporelle basse sont ensuite décomposées à nouveau selon le même procédé. Le filtrage passe bas de ces images fournit 4 nouvelles images de fréquence temporelle basse t-LL et le filtrage passe-haut de ces mêmes images fournit 4 images de fréquence temporelle haute t-LH. Le troisième niveau de décomposition fournit 2 images de fréquence temporelle basse t-LLL et 2 images de fréquence temporelle haute t-LLH. Le quatrième et dernier niveau fournit une image de fréquence temporelle basse t-LLLL et une image de fréquence temporelle haute t-LLH.

Cette décomposition temporelle est une décomposition 5 bandes temporelles qui génère donc 1 image t-LLLL, 1 image t-LLH, 2 images t-LH, 4 images t-L, et 8 images t-H par GOP de 16 images. Les images t-L, t-LL, t-LLL et bien sûr les images originales sont ignorées pour le codage en aval puisqu'elles sont à l'origine de la décomposition en sous-bandes pour fournir les images décorréées à chaque niveau. Cette décomposition permet donc une nouvelle répartition de l'énergie en générant une image utile de fréquence temporelle basse t-LLLL, qui représente une moyenne de l'ensemble du GOP et dans laquelle est concentrée l'énergie, et quatre niveaux d'images de fréquence temporelle haute de faible énergie, soit 5 bandes de fréquence. Ce sont ces images qui sont transmises au circuit d'analyse spatiale pour une décomposition spatiale en sous-bandes.

Pour réaliser les filtrages, un champ de mouvement est estimé entre chaque paire d'images à filtrer et cela pour chaque niveau. C'est la fonction de l'estimateur de mouvement 7.

D'une manière pratique, seul un champ de vecteurs mouvement est calculé, de A vers B ou de B vers A. L'autre champ de vecteurs mouvement est déduit du premier, générant des pixels non connectés, c'est à dire non affectés d'un vecteur mouvement et correspondant à des trous dans le champ de vecteurs mouvement inverse.

L'opération de filtrage est effectuée sur chaque paire d'image du GOP original, pour obtenir un premier niveau de résolution temporelle. Le

processus est répété plusieurs fois sur les images résultantes de la bande de fréquence temporelle basse, pour obtenir les niveaux de résolution temporelle suivants. Dans le cas d'un GOP de 16 images, le processus génère 8 bandes de fréquence temporelle appelées L, H, LL, LH, LLL, LLH, LLLL, LLLH.

Comme indiqué plus haut, le problème de pixels non connectés se traduit par une moins bonne décorrélation lors du filtrage temporel et donc une moins bonne compression des données. Il peut se produire à chacun des niveaux de décomposition du GOP où une estimation de mouvement est effectuée.

Selon un mode de réalisation de l'invention, un circuit de commutation du mode est mis en œuvre pour contrôler l'analyse temporelle, circuit commandé par la pertinence de l'estimation de mouvement. Il permet de gérer le cas où un nombre important de pixels non connectés est obtenu lors de l'estimation de mouvement.

La figure 4 représente un dispositif avec un tel circuit. Les mêmes références sont utilisées pour désigner les circuits semblables à ceux de la figure 1 qui ne sont donc pas décrits à nouveau.

Les informations de mouvement provenant du circuit d'estimation de mouvement référencé 2 sont transmises à un circuit de commutation de mode référencé 7. Ce circuit est relié au circuit d'analyse temporelle 1 pour lui transmettre les informations de mouvement et les informations sur le mode de codage à exploiter.

La figure 4 représente un organigramme simplifié de l'algorithme mis en œuvre pour le fonctionnement du circuit de commutation.

Soit A et B deux images successives d'un niveau de décomposition temporelle donné, disponibles à l'étape référencée 7. Le mouvement est par exemple estimé de l'image B vers l'image A et l'image A est donc l'image de référence. Une étape référencée 8 effectue une estimation de mouvement de cette image B vers l'image A. L'étape suivante 9 réalise un calcul du pourcentage P1 de pixels non connectés dans l'image de référence A par rapport au nombre de pixels de l'image, c'est à dire le nombre de pixels auxquels aucun vecteur mouvement n'est affecté. L'étape suivante référencée 10 effectue une comparaison de ce pourcentage P1 à un seuil S1 qui est une valeur prédéfinie. Cette valeur est par exemple un pourcentage de l'ordre de 20 à 25 %.

Si le nombre de pixels non connectés est inférieur ou égal à cette valeur, le champ de vecteurs mouvement est considéré comme correct et la prochaine étape est l'étape 11 qui effectue un filtrage de manière classique.

5 Les images de fréquence basse et haute sont ainsi calculées :

$$\begin{cases} H = \frac{B - MC_{A \leftarrow B}(A)}{\sqrt{2}} \\ L = \sqrt{2} \cdot A + MC_{A \leftarrow B}^{-1}(H) \end{cases}$$

Ce filtrage, équivalent à celui précédemment décrit, consiste à calculer d'abord l'image H. Cette image est obtenue par différence point à point de l'image B et de l'image A compensée en mouvement. Ainsi, on retranche à un pixel de B une certaine valeur, interpolée le cas échéant, pointée par le vecteur déplacement dans A, vecteur mouvement calculé lors de l'estimation de mouvement de l'image B vers l'image A.

L'image L est ensuite déduite à partir de l'image H et non plus de l'image B, par addition de l'image A à l'image H compensée en mouvement inverse. $MC_{A \leftarrow B}^{-1}(H)$ correspond à une « décompensation » en mouvement de l'image (H). Ainsi, on ajoute, à un pixel de A, une certaine valeur, interpolée le cas échéant, située, dans l'image H, à la base d'un vecteur déplacement de B vers A et pointant le pixel de A.

Si le pourcentage de pixels non connectés est supérieur au seuil S1, le champ de vecteurs mouvement est considéré comme non pertinent et l'étape 12 succède à l'étape 10.

S'il existe, au niveau temporel traité, une image C postérieure à l'image B, test effectué à l'étape 12, cette image C devient une nouvelle image de référence possible et le mouvement est estimé de l'image B vers l'image C, étape 13.

S'il n'y a pas d'image postérieure à l'image B au niveau courant traité, c'est l'étape 17 qui succède à l'étape 12. Cette étape code les images B et A en mode intra avec le facteur d'échelle approprié pour l'image A.

$$\begin{cases} H = B \\ L = \sqrt{2} \cdot A \end{cases}$$

30 L'étape 14, qui succède à l'étape 13, réalise un calcul de pourcentage P2 de pixels non connectés dans l'image C pour le champ de vecteurs mouvement calculé entre B et C.

L'étape suivante 15 effectue une comparaison de ce pourcentage P2 à un seuil S2 qui est une valeur prédéfinie. Cette valeur est par exemple un pourcentage de l'ordre de 20 à 25 %.

Si le nombre de pixels non connectés est inférieur ou égal à
5 cette valeur, le champ de vecteurs mouvement est considéré comme correct et la prochaine étape est l'étape 16 qui effectue un filtrage de manière classique mais avec cette image C. Il s'agit du mode « prédiction différée » (backward prediction). Les images de fréquence basse et haute sont ainsi calculées :

10

$$\begin{cases} H = \frac{B - MC_{B \rightarrow C}(C)}{\sqrt{2}} \\ L = \sqrt{2}.A \end{cases}$$

L'information de l'image A se trouve ainsi dans l'image L et l'énergie de l'image H est réduite en prenant cette référence dans le futur pour le filtrage plutôt qu'en choisissant simplement l'image B ou en la filtrant
15 à partir d'une image précédente non corrélée. On obtient ainsi une image haute fréquence de basse énergie.

Si le pourcentage de pixels non connectés est inférieur au seuil S2, l'étape 17 succède à l'étape 12. . Cette étape 17 code les images B et A en mode intra avec le facteur d'échelle approprié pour l'image A, comme
20 dans le mode MCTF ou le mode de prédiction différée précédent.

$$\begin{cases} H = B \\ L = \sqrt{2}.A \end{cases}$$

L'inverse est bien sûr possible qui consiste à prendre l'image A comme image haute fréquence (H=A) et l'image B comme image basse fréquence. On pourrait par exemple choisir, pour L, l'image possédant le
25 moins d'énergie.

Les étapes 11, 16 et 17 sont donc les modes de codage déterminés par le circuit de commutation de mode, informations transmises au circuit d'analyse temporelle.

Le choix du mode est ici effectué pour l'image complète. Il est
30 bien sûr tout aussi envisageable de faire une sélection du mode pour chacun des blocs ou macroblochs de l'image.

Le terme $\sqrt{2}$, appelé facteur d'échelle, relatif par exemple aux images L « réelles », c'est à dire obtenues par un filtrage MCTF, est dû au

filtrage MC lift. Ce facteur d'échelle, également attribué au codage en mode intra des images A, permet d'avoir la même valeur énergétique au niveau des images et vise à faciliter la mise en correspondance lors de l'estimation de mouvement pour le prochain niveau temporel, les images étant alors
 5 homogènes en termes d'énergie.

L'étape de filtrage élémentaire décrite ci-dessus est réalisée pour chaque couple d'image d'un niveau temporel donné et cela pour chaque niveau temporel jusqu'au niveau au sommet de la pyramide temporelle du GOP, quelque soit la pertinence des champs de vecteur mouvement, quelque soit le mode. Ceci permet de ne pas être affecté par de fausses
 10 coupures de scène, correspondant par exemple à un déplacement rapide de la caméra, découvrant des zones jusqu'alors cachées.

A la fin du procédé, les images A ou B qui ont été codées en mode intra à un niveau donné de décomposition et qui n'ont pas servi d'image de référence ou de prédiction pour le mode MCTF ou pour le mode de prédiction différée aux niveaux de décomposition temporelle suivants, sont remises à l'échelle pour correspondre au dernier niveau pour lequel ses pixels ont été exploités pour ces modes :

$$F_i' = F_i / \sqrt{2}^{n(F_i)}$$

20 F_i est l'image L ou H résultant du filtrage de l'image A et B et F_i' cette image F_i remise à l'échelle.

L'exposant n à appliquer au facteur d'échelle est donné en fonction du mode, par les formules suivantes :

25

- mode MCTF

$$\begin{cases} n(L) = 0 \\ n(H) = 0 \end{cases}$$
- mode Prédiction ou mode de codage intra:

$$\begin{cases} n(L) = n(A) + 1 \\ n(H) = n(B) \end{cases}$$

La valeur de n est ainsi calculée, pour une image L ou H à un niveau de décomposition donnée, en fonction de la valeur de n attribuée à l'image A ou B du niveau de décomposition précédent, utilisée
 30 respectivement pour le calcul de L ou H.

Ainsi, si les images A et B sont traitées selon le mode MCTF pour donner les images L et H, alors, quel que soit leur niveau $n(A)$ et $n(B)$, le niveau n attribué à ces images L et H est forcé à zéro.

Si le mode prédiction ou intra est utilisé pour donner les images L et H à partir des images A et B, affectées du niveau $n(A)$ et $n(B)$, le niveau de l'image L est celui $n(A)$ incrémenté de un et le niveau de l'image H est celui de l'image B, $n(B)$.

5 En se référant à la figure 2, si l'on considère une coupure de scène après la quatrième image du GOP menant à un codage en intra effectué pour la première fois au troisième niveau de décomposition pour les deux images t-LLL puis au quatrième et dernier niveau, la valeur de $n(L)$ est égale à 2 pour l'image t-LLLL et la valeur de $n(H)$ est égale à 1 pour l'image
10 t-LLH. Ainsi, en divisant deux fois l'image t-LLLL par le facteur d'échelle, on retrouve l'image t-LL qui correspond au dernier niveau auquel cette image est utilisée pour le mode MCTF.

Il est possible d'affecter un numéro à chacune des images
15 originales du GOP. Des numéros sont ensuite attribués aux images décomposées, un filtrage d'une image A numérotée s avec une image B numérotée t donnant une image de basse fréquence numérotée s, L correspondant à A, et une image de haute fréquence numérotée t, H correspondant à B. Par exemple, les numéros obtenus pour les images des
20 niveaux t-LLLL, t-LLH, t-LLH sont 1, 9, 5 et 13. Pour chaque niveau de décomposition, le calcul de n pour une image numérotée i consiste alors en une incrémentation du niveau n affecté à l'image i du niveau précédent dans le cas où cette image subit un codage intra ou prédictif pour donner une image basse fréquence L. Le compteur est inchangé dans le cas où cette
25 image subit un codage intra ou prédictif pour donner une image haute fréquence H. Le compteur est remis à zéro dans le cas où cette image subit un codage MCTF. La valeur n attribuée à chaque image originale du GOP est initialisée à zéro. Le calcul se fait à chaque niveau de décomposition jusqu'au dernier niveau.

30 En fait, la valeur n attribuée à une image décomposée correspond :

- pour une image finale obtenue par filtrage MCTF, à la valeur zéro,
- pour une image finale basse fréquence obtenue par codage
35 intra, au nombre de codages intra réalisés sur l'image de même numéro aboutissant à cette image basse fréquence, incrémenté de un,

- pour une image haute fréquence obtenue par codage intra, au nombre de codages intra successifs réalisés sur l'image de même numéro aboutissant à cette image haute fréquence,

- pour une image haute fréquence obtenue par filtrage temporel avec une image suivante, au nombre de codages intra successifs réalisés sur l'image de même numéro, incrémenté de un, aboutissant à l'image objet de ce filtrage temporel.

Cette dernière étape de remise à l'échelle vise à calibrer la dynamique des images transformées qui ne sont pas des sous-bandes temporelles «réelles», c'est à dire des images de sous-bande obtenues par filtrage temporel MCTF, en diminuant le poids de ces images.

L'image sous-bande calibrée est une image qui n'est pas exploitée pour la synthèse des deux images A et B, raison pour laquelle son coût est réduit.

- L'allocation de débit est réalisée par le paquetiseur 5. L'orthogonalité de la transformée temporelle, et également sa normalisation, sont des points essentiels pour que l'allocation de débit soit optimale parmi les différentes images. Ce qui est réalisé de façon naturelle lors d'un filtrage temporel MCTF normalisé pourrait devenir un problème lorsque d'autres modes tels que le mode de prédiction différée ou le mode de codage en intra sont mis en œuvre. Si, dans une étape de codage, une image de fréquence basse est obtenue, en mode intra ou prédictif, en multipliant l'image A par un facteur d'échelle pour atteindre la dynamique du niveau temporel, c'est à dire pour d'obtenir une image de même énergie que les images obtenues par filtrage MCTF, cette pondération n'est plus souhaitable au niveau de l'allocation de bits. Au contraire, un soin particulier doit être pris pour favoriser les images de fréquence basse «réelles» obtenues par filtrage MCTF, les images de référence pertinentes, par exemple en donnant moins de poids aux images codées en mode intra ou prédictif. En effet, les images de fréquence basse réelles, comme on va le voir ci-après, sont exploitées au décodeur pour la reconstruction des deux images A et B au niveau inférieur de décomposition, selon la pyramide hiérarchique, B étant fonction de A, ce qui n'est pas le cas pour les images codées en mode intra ou prédictif.

L'invention concerne également un décodeur exploitant les données codées selon le procédé précédemment décrit. La figure 5 représente un tel décodeur.

Le flux binaire correspondant aux images codées selon le procédé précédemment décrit sont transmises à l'entrée du décodeur. Elles sont transmises en parallèle à l'entrée d'un circuit de décodage entropique 18 et à l'entrée d'un circuit de décodage du mouvement 19. Le circuit de décodage entropique est relié successivement à un circuit de synthèse spatiale 20 et à un circuit de synthèse temporelle 21. Ce dernier circuit reçoit des informations de mouvement du circuit de décodage du mouvement. La sortie du circuit de synthèse temporelle est reliée à l'entrée d'un circuit de post-filtrage 22 dont la sortie est la sortie du décodeur.

Le train binaire est donc traité par le circuit de décodage entropique 18 qui réalise les opérations inverses du circuit de codage entropique au codeur et décode les coefficients ondelettes spatio-temporels et les modes de filtrage.

Le circuit de décodage du mouvement 19 récupère dans le flux et décode les informations relatives au mouvement pour transmettre au circuit 21 les champs de mouvement nécessaires à la synthèse temporelle.

Le circuit de synthèse spatiale 20 reconstruit les images correspondant aux différentes sous-bandes temporelles. Les images sous-bandes reconstruites sont ensuite mises à l'échelle à partir du facteur d'échelle approprié, pour effectuer la synthèse temporelle. Cette mise à l'échelle se fait en fonction des informations de mode transmises et relatives aux images utilisées lors de la décomposition du GOP original, au codeur, permettant d'attribuer les valeurs de n aux différentes images.

$$F_i = F_i' \cdot \sqrt{2}^{n(F_i)}$$

Ensuite, le circuit de synthèse temporelle 21 reconstruit les images transformées jusqu'au niveau temporel souhaité, selon leur mode :

30

- mode MCTF

$$\begin{cases} A = \frac{L - MC_{A \leftarrow B}^{-1}(H)}{\sqrt{2}} \\ B = \sqrt{2} \cdot H + MC_{A \leftarrow B}(A) \end{cases}$$

- mode de prédiction backward

$$\begin{cases} A = L/\sqrt{2} \\ B = \sqrt{2}.H + MC_{B \rightarrow C}(C) \end{cases}$$

- mode de codage intra

$$\begin{cases} A = L/\sqrt{2} \\ B = H \end{cases}$$

5 Les informations de mouvement nécessaires à la compensation de mouvement proviennent du circuit de décodage du mouvement. Le circuit de synthèse temporelle réalise un décodage en fonction de l'information de mode de codage affectée à l'image à décoder.

10 Les images en sortie du circuit 21 sont donc reconstruites par synthèse temporelle à partir des coefficients ondelettes temporels. Une étape finale de post-traitement est appliquée par le circuit de post filtrage 22 en réalisant un filtrage sur l'image permettant de réduire les artefacts de type effets de bloc.

15 Le pourcentage P1 et/ou P2 utilisé pour déterminer le mode de commutation est choisi de l'ordre de 20 à 25%. Ce pourcentage a été obtenu de manière empirique et d'autres valeurs peuvent bien sûr être exploitées pour la mise en œuvre du procédé de commutation.

20 Le mouvement exploité pour le filtrage temporel d'une paire d'images peut être obtenu par simplification ou élaguage, plus connu sous l'appellation anglaise de « pruning », du champ de vecteurs mouvement calculé par l'estimation de mouvement, permettant de limiter le coût de codage du mouvement.

25 Une variante de l'invention consiste à combiner une structure de GOP adaptatif au procédé précédemment décrit. La taille du GOP est alors variable, dépendant par exemple de paramètres tels que le mouvement dans la séquence.

L'invention s'applique aussi bien à un circuit d'analyse temporelle de type hybride que de type codage sous-bande.

30 Les applications de l'invention concernent la compression vidéo avec prédiction temporelle.

REVENDECATIONS

1 Procédé de codage d'une séquence d'images comprenant une analyse temporelle hiérarchique (1) d'un groupe d'images réalisant un filtrage
5 temporel compensé en mouvement de paires successives d'images (11) pour fournir des images de fréquence temporelle basse et des images de fréquence temporelle haute à différents niveaux de décomposition temporelle, cette analyse réalisant, pour un niveau de décomposition temporelle donné et pour une paire d'images de fréquence temporelle basse,
10 une étape d'estimation de mouvement (8) d'une image courante B vers une image précédente de référence A pour fournir des vecteurs mouvement puis un filtrage temporel compensé en mouvement (11) de ces images pour fournir une image de fréquence temporelle basse (L) et une image de fréquence temporelle haute (H) d'un niveau de décomposition supérieur, le
15 dit filtrage temporel étant remplacé par un codage en mode intra (16, 17) pour obtenir au moins une image de fréquence basse (L) ou haute (H) si l'image courante a un niveau de corrélation avec une image précédente inférieur à un seuil (10, 15), les images obtenues de fréquence basse (L) étant alors mises à l'échelle pour être adaptées, au niveau énergétique, aux
20 images obtenues par le dit filtrage temporel compensé en mouvement, caractérisé en ce que, parmi l'image basse fréquence et les images haute fréquence finales décomposées obtenues en fin d'analyse :

- il sélectionne les images obtenues par un codage intra d'une image du niveau de décomposition inférieur avec la condition
25 supplémentaire, pour les images haute fréquence, que cette image provient elle-même d'un codage intra.

- il calibre l'image sélectionnée en effectuant au moins une étape inverse de l'étape de mise à l'échelle.

30 2 Procédé selon la revendication 1, caractérisé en ce que le nombre d'étape inverses effectuées correspond au nombre de codages intra successifs d'une image basse fréquence (L) pour arriver à l'image sélectionnée s'il s'agit d'une image sélectionnée haute fréquence, ce nombre

étant incrémenté de un s'il s'agit de l'image sélectionnée basse fréquence (L).

3 Procédé selon la revendication 2, caractérisé en ce qu'il comporte, pour le calcul d'une image fréquence basse L ou haute H, un
5 filtrage temporel entre l'image courante et une image suivante (16) de la paire d'images suivante, si la corrélation entre l'image courante et l'image précédente est inférieure à un seuil (10) et si la corrélation entre l'image courante et cette image suivante est supérieure à un seuil (15), l'autre image H ou L étant obtenue par codage intra (16) et en ce que cette opération de
10 filtrage est assimilée au codage intra et non pas au filtrage temporel pour l'étape de sélection.

4 Procédé selon la revendication 3, caractérisé en ce qu'il affecte à chaque image du groupe d'images un numéro d'image, en ce qu'il effectue
15 un suivi de ces images numérotées lors de la décomposition en attribuant un compteur pour chaque numéro, ce compteur étant actualisé à chaque étape :

- le compteur est incrémenté chaque fois qu'une image de fréquence basse (L) est obtenue en mode intra,
- le compteur est inchangé chaque fois qu'une image de
20 fréquence haute (H) est obtenue en mode intra ou lors d'un filtrage temporel avec une image suivante,
- le compteur est remis à zéro chaque fois qu'une image est obtenue par filtrage temporel compensé en mouvement, avec une image précédente.

25

5 Procédé selon la revendication 1, caractérisé en ce que les images de fréquence haute H et de fréquence basse L sont obtenues, lors du filtrage temporel compensé en mouvement de deux images successives A et B à partir des opérations suivantes :

30

$$\begin{cases} A = \frac{L - MC_{A \leftarrow B}^{-1}(H)}{\sqrt{2}} \\ B = \sqrt{2}.H + MC_{A \leftarrow B}(A) \end{cases}$$

MC correspondant à la compensation en mouvement selon le champ de vecteurs de B vers A, de l'image H.

- 5 6 Procédé selon la revendication 5, caractérisé en ce que les images L et H sont obtenues, par codage intra, selon les formules

$$\begin{cases} H = B \\ L = \sqrt{2}.A \end{cases}$$

et en ce que les images H et L sont obtenues, par filtrage avec l'image suivante pour H et par codage intra pour L, selon les formules suivantes :

10
$$\begin{cases} H = \frac{B - MC_{B \rightarrow C}(C)}{\sqrt{2}} \\ L = \sqrt{2}.A \end{cases}$$

MC correspondant à la compensation en mouvement selon le champ de vecteurs de B vers C, de l'image C.

- 15 7 Procédé selon la revendication 1, caractérisé en ce que les images calibrées obtenues par analyse temporelle (1) sont ensuite traitées par analyse spatiale (3).

- 20 8 Procédé selon la revendication 1, caractérisé en ce que le niveau de corrélation est calculé en prenant en compte le nombre de pixels connectés, c'est à dire reliés par un vecteur de mouvement.

- 25 9 Procédé de décodage d'une séquence d'images codées selon le procédé de la revendication 1, caractérisé en ce qu'il effectue une étape de calibrage inverse (21) d'images sélectionnées à décoder, la sélection des images et le nombre d'étapes inverses étant fonction d'une information associée à l'image à décoder.

10 Procédé de décodage d'une séquence d'images codées selon
la procédé de la revendication 4, caractérisé en ce qu'il effectue une étape
de calibrage inverse (21) d'images sélectionnées à décoder, la sélection et le
nombre d'étapes inverses étant fonction de la valeur d'un compteur affecté à
5 l'image lors du codage.

11 Codeur pour la mise en œuvre du procédé selon la
revendication 3, comportant un circuit d'analyse temporelle (1) exploitant le
filtrage temporel compensé en mouvement et le codage intra, caractérisé en
10 ce que le circuit sélectionne, parmi l'image basse fréquence et les images
haute fréquence finales décomposées obtenues en fin d'analyse, les images
obtenues par un codage intra d'une image du niveau de décomposition
inférieur ou par un filtrage temporel entre l'image d'une paire d'images et
l'image suivante de la paire d'images suivante du niveau de décomposition
15 inférieur, avec la condition supplémentaire, pour les images haute fréquence,
que cette image provient elle-même d'un codage intra et en ce qu'il effectue
au moins une étape inverse de l'étape de mise à l'échelle pour les images
sélectionnées.

20 12 Décodeur pour le décodage de séquence d'images codées
selon le procédé de la revendication 1, comportant un circuit de synthèse
temporelle (21), caractérisé en ce que le circuit comporte des moyens pour
effectuer un calibrage inverse d'images à décoder, la sélection des images et
le nombre de calibrages inverses étant fonction d'une information associée à
25 l'image à décoder et reçue par le décodeur.

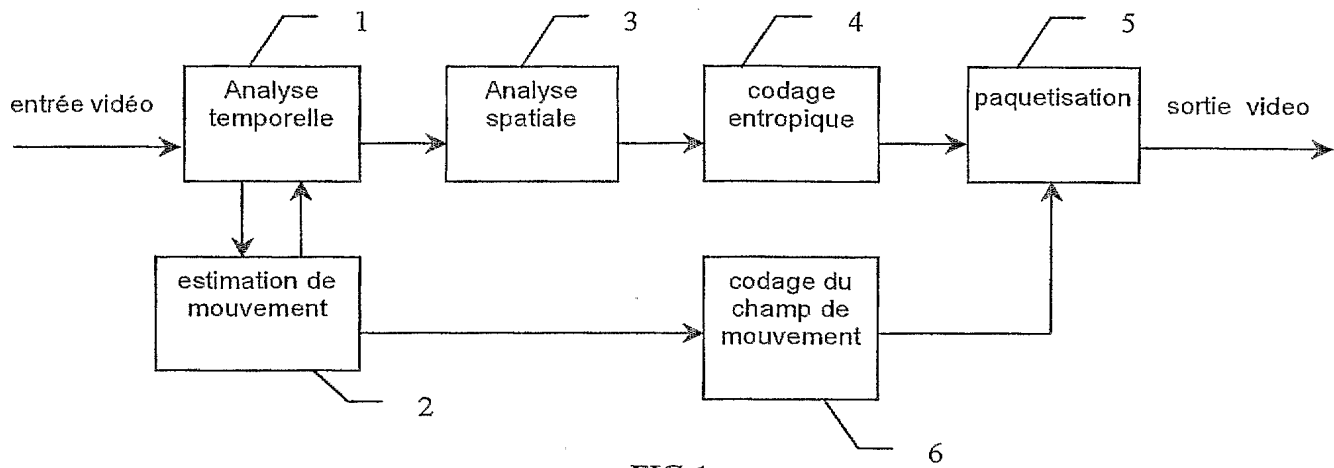


FIG.1

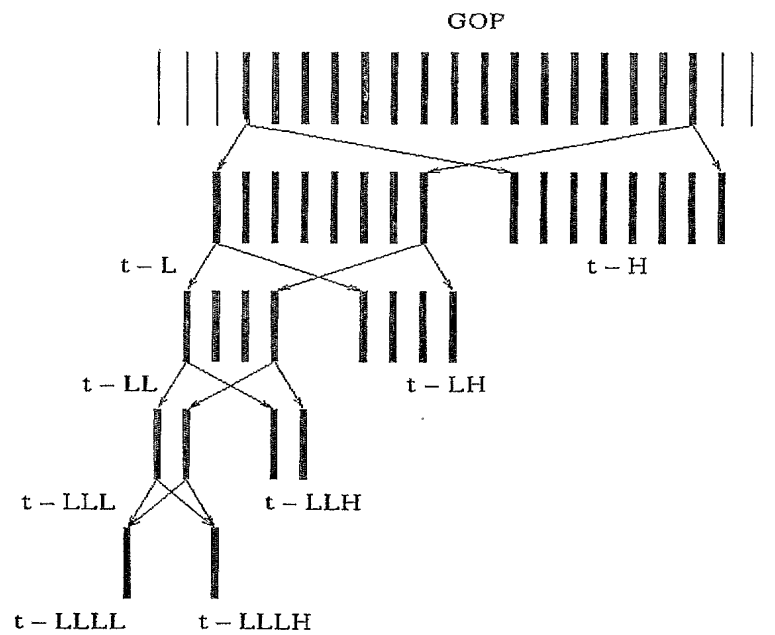


FIG.2

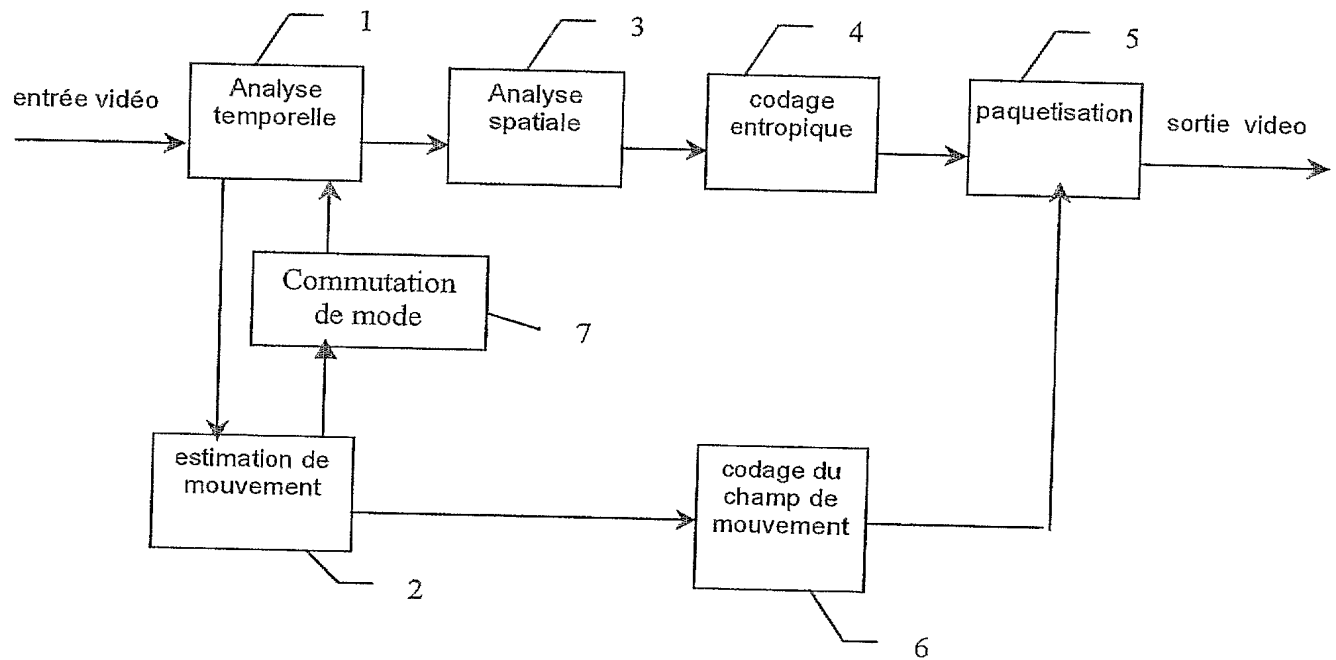


FIG.3

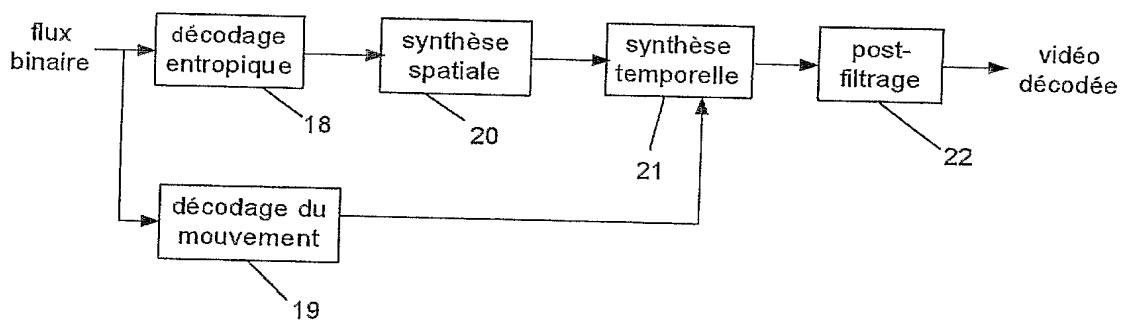


FIG.5

3/3

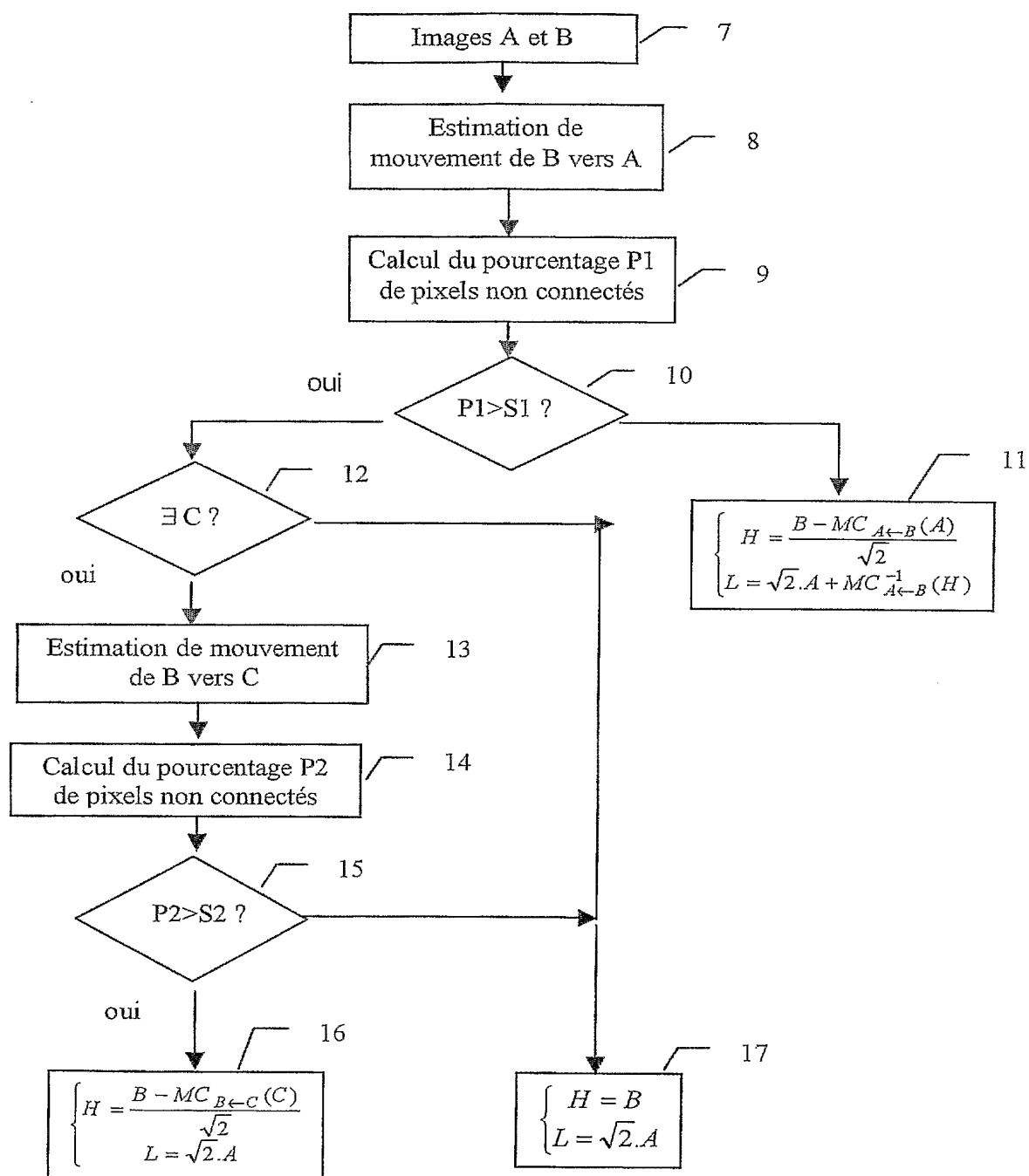


FIG.4



BREVET D'INVENTION CERTIFICAT D'UTILITE

Désignation de l'inventeur

Vos références pour ce dossier	PF040033
N°D'ENREGISTREMENT NATIONAL	
TITRE DE L'INVENTION	
	PROCEDE DE CODAGE D'UNE SEQUENCE D'IMAGES PAR ANALYSE TEMPORELLE HIERARCHIQUE
LE(S) DEMANDEUR(S) OU LE(S) MANDATAIRE(S):	
DESIGNE(NT) EN TANT QU'INVENTEUR(S):	
Inventeur 1	
Nom	FRANCOIS
Prénoms	Edouard
Rue	18 Allée du Locar
Code postal et ville	35890 BOURG-DES-COMPTES
Société d'appartenance	THOMSON
Inventeur 2	
Nom	BOISSON
Prénoms	Guillaume
Rue	12 rue Jean Malo-Renault
Code postal et ville	35000 RENNES
Société d'appartenance	THOMSON
Inventeur 3	
Nom	VIERON
Prénoms	Jérôme
Rue	14 Square Louis Massignon
Code postal et ville	35000 RENNES
Société d'appartenance	THOMSON
Inventeur 4	
Nom	ROBERT
Prénoms	Philippe
Rue	7 Allée du Bois Louët
Code postal et ville	35235 THORIGNE-FOUILLARD
Société d'appartenance	THOMSON
Inventeur 5	
Nom	MARQUANT
Prénoms	Gwenaëlle
Rue	1 rue du Général de Gaulle
Code postal et ville	35340 LIFFRE
Société d'appartenance	THOMSON

La loi n°78-17 du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire. Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI.

Signé par

Signataire: FR, THOMSON, B.Ruellan

Emetteur du certificat: DE, D-Trust GmbH, D-Trust for EPO 2.0

Fonction

Mandataire agréé (Mandataire 1)

FC /FR 05 50109

